

Figure 1.

```
Sequence type explicitly set to Protein
Sequence format is Pearson
Sequence 1: new_S100_cytokine      104 aa
Sequence 2: G491246                110 aa
Sequence 3: W27152                 98 aa
Start of Pairwise alignments
Aligning...
Sequences (1:2) Aligned. Score: 23
Sequences (1:3) Aligned. Score: 34
Sequences (2:3) Aligned. Score: 29
Start of Multiple Alignment
There are 2 groups
Aligning...
Group 1:                          Delayed
Group 2:                          Delayed
Sequence:1      Score:0
Sequence:3      Score:839
Sequence:2      Score:724
Alignment Score 444
CLUSTAL-Alignment file created [/data4/genetools/lrastelli4630clustalw]
```

Multiple Alignment:

```
new_S100_cytokine  MGQCRSANAEDQEFSEYERAEITGKNEHOYSWEGS-ETITLTPSELRLVTQQLPHLNP
W27152            -----MAEPISELESIEETVITIFFTEARQESRRKISLSVNERHELVTDQLPHLLP
G491246            -----MCLEENIETLNTTFHOYSWKLEHPDTLNQGEFHELVKDKLQNFLP

new_S100_cytokine  SNGC----LEEHYANLQSCNDSKEERSEWEIIGEAANSVKLERP....VRSH...(SEQ ID NO:6)
W27152            DVGS----LDEHHKSLVWQDSSELKENEYWRIGELAKEIRKED....LRIRKK...(SEQ ID NO:10)
G491246            KENKNEKVIEHIMEDLTHADKQLSEEFIMLMARLTWASHEAMHEGDEGPEHHHKPGLG

new_S100_cytokine  ----
W27152            ----
G491246            EGTP (SEQ ID NO:11)
```

Figure 2.

Multiple Alignment:

```
new_S100_cytokine      -----DNRTLKCPDTWS-TMGQCRSANAEDAQEFSDVERAIETLIKNFHQYS
7971c.7__r0s0-212.2__EXT SISSCGAGYRTDDKTQLTEERTSWPBTMGQCRSANAEDAQEFSDVERAIETLIKNFHKYS

new_S100_cytokine      WEFGKETLTPS ELRDLVTQQLPHLMPSNCGLEEKIANLGSCNDSKLEFRSFWELIGEAAH
7971c.7__r0s0-212.2__EXT WAKKETLTPEL RDLVTQQLPHLMPSNCGLEEKIANLGNENDSKLEFCFWELIGEAAH

new_S100_cytokine      SVKMERPVGRH  (SEQ ID NO:6)
7971c.7__r0s0-212.2__EXT SVKMERPVTRS  (SEQ ID NO:3)
```

Figure 3.

hmmpfam - search a single seq against HMM database

HMMER is freely distributed under the GNU General Public License (GPL).

HMM file: pfamHMMs

Sequence file: /data4/genetools/lrastelli4423Aa315020ProteinFasta.txt

Query: AA315020

Scores for sequence family classification (score includes all domains):

Model	Description	Score	E-value	N
S_100	S-100/ICaBP type calcium binding domain	40.9	2.9e-08	1

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
S_100	1/1	32	74 ..	1	44 []	40.9	2.9e-08

Alignments of top-scoring domains:

S_100: domain 1 of 1, from 32 to 74: score 40.9, E = 2.9e-08

```

      *-->LEKkaietiInvFhqYSgreGdkdtLsKkELKellekELpnflkn<--*
          K+aiet+I+ FhqYS eG k tL+ EL+ L++++Lp+ +
AA315020  32  VERAIETLIKRFHQYS-VEGCKETLTPSELRLDLVTQQLPHLMPS  74(SEQ ID NO:33)

```

//

Figure 4B

Table 3
AY007220
Consensus

10	20	30	40	50	60
.....					
MGQCRSANAEDAQEFSDVERAIETLIKNFHRYSVASKKETLTPELRDLVTQQLPHLMPS					
MGQCRSANAEDAQEFSDVERAIETLIKNFHRYSVEGCKETLTPELRDLVTQQLPHLMPS					
MGQCRSANAEDAQEFSDVERAIETLIKNFHYSVSKETLTPELRDLVTQQLPHLMPS					

Table 3
AY007220
Consensus

70	80	90	100
.....			
NCGLEEKIANLGNCNDSKLEFCSFWELIGEAAKSVKLERPV (SEQ ID NO:3)			
NCGLEEKIANLGSCNDSKLEFRSFWELIGEAAKSVKLERPV (SEQ ID NO:39)			
NCGLEEKIANLGNCNDSKLEFSFWELIGEAAKSVKLERPV (SEQ ID NO:40)			

Figure 4A.

BLOCKS Protein Domain Analysis of new_S100_cytokine

Probe Size: 104 Amino Acids

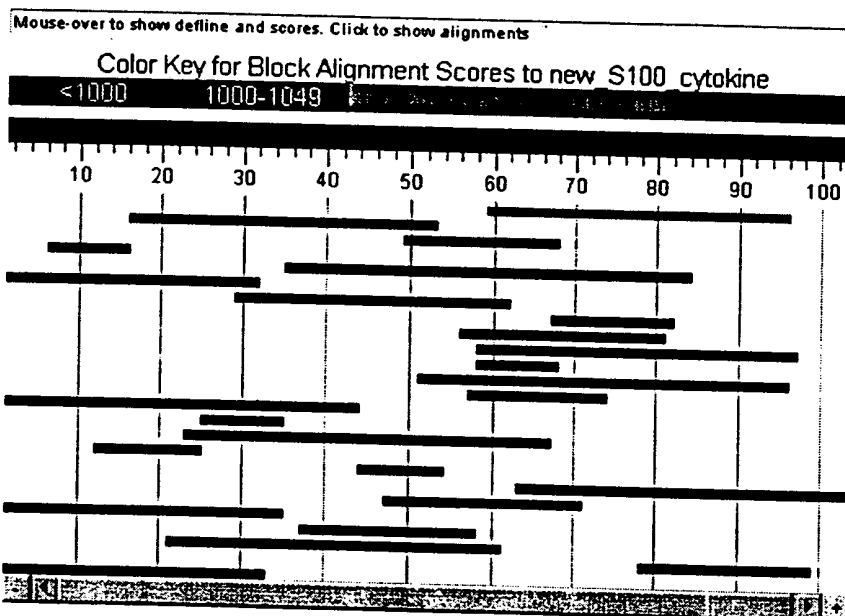
Probe File: lrastelliblocks.seq

Target File (s) : blocks.dat

Records Searched: 4034

Scores Done: 4034

Alignments Done: 535470



Accession	Description	Strength	Score	HT	Accession	Score	HT
HL00202B	3-100/ICAP type calcium binding protein.	1236	1057	0	59	HL00202B	1057
HL00202A	3-100/ICAP type calcium binding protein.	1245	1038	0	15	HL00202A	1038
HL00274A	Bacterial type II secretion system protein F	1456	1011	0	49	HL00274A	1011
HL00377B	Ubiquitin carboxyl-terminal hydrolase family	1227	991	0	49	HL00377B	991
HL00524C	Bacterial chemotaxis sensory transducers prot	1750	990	0	6	HL00524C	990
HL00527I	Phosphoenolpyruvate carboxylase (ATP) prote	1411	989	0	25	HL00527I	989
HL00704A	Prokaryotic-type carboxic anhydrases; proteins	1529	987	0	0	HL00704A	987
HL01017E	Ergosterol biosynthesis ERG4/ERG14 family pr	1499	983	0	29	HL01017E	983
HL00310E	Lysosome-associated membrane glycoproteins du	1632	976	0	67	HL00310E	976
HL00429C	Phosphofructokinase proteins.	1581	976	0	56	HL00429C	976
HL00007C	PH domain proteins profile.	990	976	0	58	HL00007C	976
HL00459	Myotoxin proteins.	1273	974	0	56	HL00459	974
HL00007C	Phosphatidylinositol-specific phospholipase X	1493	974	0	58	HL00007C	974
HL01077B	Glypican proteins.	1508	972	0	61	HL01077B	972
HL01077B	Membrane attack complex components / perforin	1127	972	0	57	HL01077B	972
HL01120D	Urease nickel ligands proteins.	1692	970	0	-1	HL01120D	970
HL01762B	Phosphoglycerate mutase family phosphohistidi	1290	968	0	23	HL01762B	968
HL00879A	Ribosomal protein L13 proteins.	1092	967	0	11	HL00879A	967
HL00822B	2'-5'-oligoadenylate synthetases proteins.	1826	967	0	44	HL00822B	967
HL01005C	Formate and nitrite transporters proteins.	1552	964	0	63	HL01005C	964
HL00161A	Glycoprotein hormones: beta chain proteins.	1528	963	0	12	HL00161A	963
HL00662A	Vinculin family tail-binding region proteins	1567	963	0	1	HL00662A	963
HL00743A	Beta-lactamases class B proteins.	1500	960	0	37	HL00743A	960
HL01021B	Heat shock hsp61 proteins family profile.	1200	960	0	21	HL01021B	960
HL01062C	Hydroxymethylglutaryl-coenzyme A lyase protei	1900	960	0	70	HL01062C	960

Figure 4B.

Table 3
AA007220
Consensus

```

      10      20      30      40      50      60
.....|.....|.....|.....|.....|.....|
MGQCRSANAEDAQEFSDVERAIETLIKNNFHYSVAGKKEITLTPSELRLDVTQQLPHLMPS
MGQCRSANAEDAQEFSDVERAIETLIKNNFHYSVAGKKEITLTPSELRLDVTQQLPHLMPS
MGQCRSANAEDAQEFSDVERAIETLIKNNFHYSVAGKKEITLTPSELRLDVTQQLPHLMPS

```

Table 3
AA007220
Consensus

```

      70      80      90      100
.....|.....|.....|.....|
NCGLEEKIANLGNCNDSKLEFRSFWELIGEAAKSVKLERPV (SEQ ID NO:3)
NCGLEEKIANLGNCNDSKLEFRSFWELIGEAAKSVKLERPV (SEQ ID NO:39)
NCGLEEKIANLGNCNDSKLEFRSFWELIGEAAKSVKLERPV (SEQ ID NO:40)

```

Figure 4C.

Table 6
AA007220
Consensus

```

      10      20      30      40      50      60
.....|.....|.....|.....|.....|.....|
MGQCRSANAEDAQEFSDVERAIETLIKNNFHOYSVEGGKETLTPSELRLDVTQQLPHLMPS
MGQCRSANAEDAQEFSDVERAIETLIKNNFHOYSVEGGKETLTPSELRLDVTQQLPHLMPS
MGQCRSANAEDAQEFSDVERAIETLIKNNFHOYSVEGGKETLTPSELRLDVTQQLPHLMPS

```

Table 6
AA007220
Consensus

```

      70      80      90      100
.....|.....|.....|.....|
NCGLEEKIANLGSCNDSKLEFRSFWELIGEAAKSVKLERPVRGH (SEQ ID NO:6)
NCGLEEKIANLGSCNDSKLEFRSFWELIGEAAKSVKLERPVRGH (SEQ ID NO:39)
NCGLEEKIANLGSCNDSKLEFRSFWELIGEAAKSVKLERPVRGH (SEQ ID NO:40)

```

Figure 4D.

Table 3
gi|4139958|pdb|1MHO|
PROTEIN MRP-126
ICTACALCIN
CALGRANULIN B
Consensus

```

      10      20      30      40
.....|.....|.....|.....|
EKALETLIKNNFHYSV-AGKKEITLTPSELRLDVTQQLPHLM (SEQ ID NO:3)
EKAIVVALIDVFHOYSGREGDTHKLEKSELRLDVTQQLPHLM (SEQ ID NO:41)
EKAIVVALIDVFHOYSRREGDTHKLEKSELRLDVTQQLPHLM (SEQ ID NO:42)
EKAIVVALIDVFHOYSRREGDTHKLEKSELRLDVTQQLPHLM (SEQ ID NO:43)
EKAIVVALIDVFHOYSRREGDTHKLEKSELRLDVTQQLPHLM (SEQ ID NO:44)
EKAIVVALIDVFHOYSRREGDTHKLEKSELRLDVTQQLPHLM (SEQ ID NO:45)

```

Figure 4E.

Table 6
gi|4139958|pdb|1MHO|
PROTEIN MRP-126
CALGRANULIN B
CALGRANULIN B
Consensus

```

      10      20      30      40
.....|.....|.....|.....|
EKALETLIKNNFHOYSV-EGGKETLTPSELRLDVTQQLPHLM (SEQ ID NO:6)
EKAIVVALIDVFHOYSGREGDTHKLEKSELRLDVTQQLPHLM (SEQ ID NO:41)
EKAIVVALIDVFHOYSRREGDTHKLEKSELRLDVTQQLPHLM (SEQ ID NO:42)
EKAIVVALIDVFHOYSRREGDTHKLEKSELRLDVTQQLPHLM (SEQ ID NO:44)
EKAIVVALIDVFHOYSRREGDTHKLEKSELRLDVTQQLPHLM (SEQ ID NO:46)
EKAIVVALIDVFHOYSRREGDTHKLEKSELRLDVTQQLPHLM (SEQ ID NO:47)

```

Figure 5

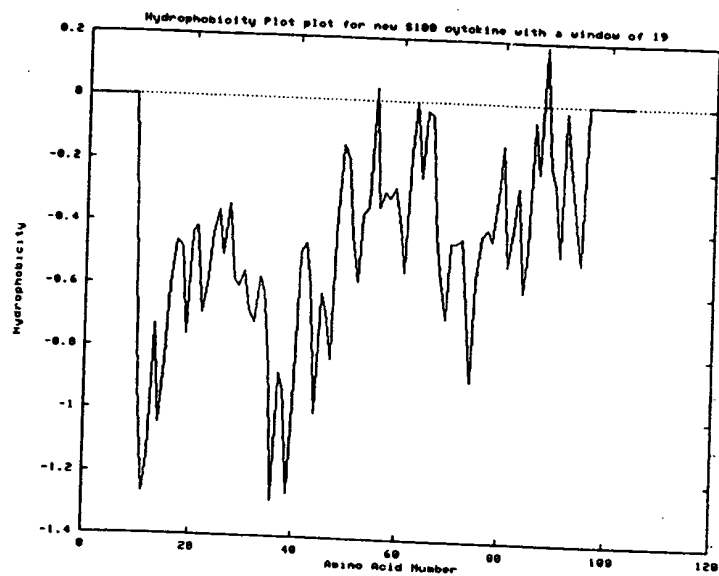


Figure 6

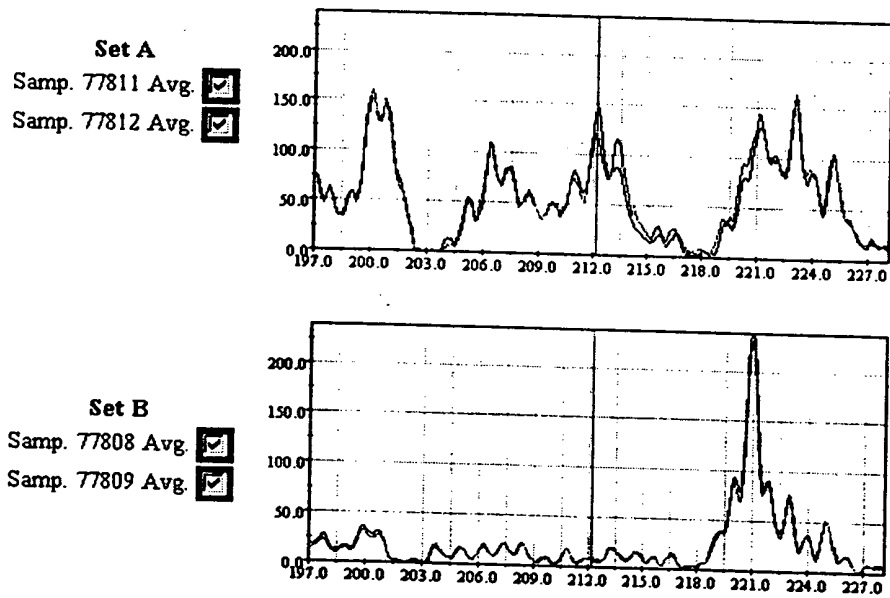


Figure 7

***** Contig 1 *****	
65677221+	GAATTCCAGAGGGAGTTCTCAGTGCCCCCGGACAGGCCTCTCCAGCTTCACACTCTTGGC
AA315020-	TGCCCCCGGACAGTCCTCTCNAGCTTCACACTCTTGGC
consensus	GAATTCCAGAGGGAGTTCTCAGTGCCCCCGGACAGGCCTCTCCAGCTTCACACTCTTGGC
65677221+	CGCTTCTCCAATCAGCTCCCAGAAACTCCTGAACTCCAGTTTAGAGTCATTGCAGCTGCC
AA315020-	CGCTTCTCCAATCAGCTCCCAGAAACTCCTGAACTCCAGTTTAGAGTCATTGCAGCTGCC
consensus	CGCTTCTCCAATCAGCTCCCAGAAACTCCTGAACTCCAGTTTAGAGTCATTGCAGCTGCC
65677221+	CAGGTTGGCAATTTTCTCTTCCAGGCCACAGTTGCTCGGCATGAGATGGGGCAGCTGCTG
AA315020-	CAGGTTGGCAATTTTCTCTTCCAGGCCANAGTTGCTCGGCATGAGATGGGGCAGCTGCTG
consensus	CAGGTTGGCAATTTTCTCTTCCAGGCCACAGTTGCTCGGCATGAGATGGGGCAGCTGCTG
65677221+	GGTGACCAGGTCCCGTAGCTCAGAAGGGGTCAGCGTCTCCTTCCCACCCTCCACGGAGTA
AA315020-	GGTGACCAGGTCCCGTAGCTCAGAAGGGGTCAGCGTCTCCTTCCCACCCTCCACGGAGTA
consensus	GGTGACCAGGTCCCGTAGCTCAGAAGGGGTCAGCGTCTCCTTCCCACCCTCCACGGAGTA
65677221+	CTGGTGAAAGTTCTTGATGAGGGTCTCAATGGCCCTCTCCACATCACTGAATTC (SEQ ID NO:37)
AA315020-	CTGGTGAAAGTTCTTGATGAGGGTCTCAATGGCCCTCTCCACATCACTGAATTCCTGAGC
consensus	CTGGTGAAAGTTCTTGATGAGGGTCTCAATGGCCCTCTCCACATCACTGAATTCCTGAGC
AA315020-	ATCCTCTGCGTTGGCTGACCGACACTGTCCCATGGTGCTCACTGTGTCTGGTCCTTTGGT
consensus	ATCCTCTGCGTTGGCTGACCGACACTGTCCCATGGTGCTCACTGTGTCTGGTCCTTTGGT
AA315020-	GAGAGTTCTGTTGTCCTAT (SEQ ID NO:4)
consensus	GAGAGTTCTGTTGTCCTAT (SEQ ID NO:5)